

mediaWalker II : ウェブコンテンツと連携した ニュース映像アーカイブ閲覧インタフェース

井手 一郎^{†,††} 高橋 友和^{†††} 木下 智義^{††††} 奥岡 知樹^{†,††††} 出口 大輔[†]
佐藤 真一^{††} 村瀬 洋[†]

† 名古屋大学 大学院情報科学研究科 〒464-8601 名古屋市千種区不老町1
†† 情報・システム研究機構 国立情報学研究所 〒101-8430 東京都千代田区一ツ橋 2-1-2
††† 岐阜聖徳学園大学 経済情報学部 〒500-8288 岐阜市中鶉 1-38
†††† (株) ネットコンパス 〒104-0033 東京都中央区新川 2-14-4 510ビル6階
††††† 現在, エヌ・ティ・ティ・コミュニケーションズ(株)
E-mail: † {ide,ddeguchi,murase}@is.nagoya-u.ac.jp, okuoka@murase.m.is.nagoya-u.ac.jp,
†† {ide,satoh}@nii.ac.jp, ††† ttakahashi@gifu.shotoku.ac.jp, †††† kino@netcompass.co.jp

あらまし 筆者らは, 大規模なニュース映像アーカイブ中の映像を, 時系列意味構造「トピックスレッド」に沿ってストーリー単位で閲覧できるインタフェースの開発に取り組んでいる. 本デモンストレーションでは, 新たに, 各ストーリーに関連する様々なウェブコンテンツとともに映像を閲覧できるインタフェース「mediaWalker II」を紹介する. キーワード 映像アーカイブ, ウェブコンテンツ, ニュース映像, トピックスレッド, 閲覧インタフェース

mediaWalker II: A News Video Archive Browsing Interface Associated with Web Contents

Ichiro IDE^{†,††}, Tomokazu TAKAHASHI^{†††}, Tomoyoshi KINOSHITA^{††††}, Tomoki OKUOKA^{†,††††},
Daisuke DEGUCHI[†], Shin'ichi SATOH^{††}, and Hiroshi MURASE[†]

† Nagoya University, Graduate School of Information Science
1 Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan
†† Research Organization of Information and Systems, National Institute of Informatics
2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan
††† Gifu Shotoku Gakuen University, Faculty of Economics and Information
1-38 Naka-Uzura, Gifu 500-8288, Japan
†††† NetCOMPASS Ltd. 6F, 510 Bldg., 2-14-4 Shinkawa, Chuo-ku, Tokyo 104-0033, Japan
††††† Currently at NTT Communications, Corp.
E-mail: † {ide,ddeguchi,murase}@is.nagoya-u.ac.jp, okuoka@murase.m.is.nagoya-u.ac.jp,
†† {ide,satoh}@nii.ac.jp, ††† ttakahashi@gifu.shotoku.ac.jp, †††† kino@netcompass.co.jp

Abstract We have been developing an interface that allows browsing footages in a large-scale news video archive per story based on a chronological semantic structure “Topic thread”. In this demonstration, we will introduce a new interface “mediaWalker II”, which allows browsing footages along with Web contents related to each story.

Key words Video archive, Web contents, news video, topic thread, browsing interface

1. はじめに

近年, 記憶装置の大容量化に伴い, 大量の映像を蓄積できるようになった. 我々は2001年以来, 特定のニュース番組の映像を蓄積してきた. これらの映像の長さは, のべ1,600時間にもなるため, 何らかのトピックに沿

った映像を手で検索しながら追跡することは難しい. そこで我々は, 大規模ニュース映像アーカイブ中の映像を, 時系列意味構造「トピックスレッド」に沿ってストーリー単位で閲覧するインタフェースの開発に取り組んできた[3]. 本デモンストレーションでは, これまでに開発した mediaWalker インタフェース [2], [4] に対して, 新たに

各ストーリーに関連するウェブコンテンツを閲覧できるようにした mediaWalker II インタフェースを紹介する。

なお、トピックスレッド構造の構築方法及び mediaWalker インタフェースの基本機能については、文献 [2] ~ [4] にゆずり、以下ではニュース映像中の各ストーリーとウェブコンテンツとの関連付け方法について簡単に紹介する。

2. ウェブコンテンツとの関連付け

2.1 Wikipedia エントリとの関連付け

まず、オンライン百科事典 Wikipedia^(注1)のエントリとの関連付けを行なう。これにより、Wikipedia エントリ自体が関連付くとともに、エントリ名をニュースストーリーの索引とみなして、他のウェブコンテンツを関連付けの際の検索に利用する。

以下に、関連付け方法を簡単に紹介する。対象とする全 Wikipedia エントリに以下の処理を施すと、各ニュースストーリーに対して、複数のエントリが関連付くこともある。詳細は文献 [5] を参照されたい。

(1) ニュース映像と Wikipedia エントリとの関連付け

Wikipedia 記事から日付に関する表現を全て抽出する。ニュース映像が放送された日かその前日の日付が含まれる場合、そのエントリを照合対象とする。これにより関連付けに要する計算時間を短縮する。

次に、ニュースストーリーの文字放送字幕 (Closed-caption) と Wikipedia 記事の各々の名詞頻度ベクトルの余弦距離が一定以下のものを関連付ける。

(2) トピックスレッド構造に沿った関連付けの拡張

毎日放送されるニュース映像と比べると、Wikipedia 記事には綿密に経緯が記述されているとは限らないため、

(1) の処理で照合対象からもれるエントリがある。そこで、事前に構築しておいたトピックスレッド構造に沿って、当該ニュースストーリーの前後のストーリーについても照合する。照合の結果関連付くものがなくなるまで、トピックスレッド構造をたどって関連付けを拡張する。

2.2 一般のウェブコンテンツとの関連付け

各ニュースストーリーにつき、その放送日と、2.1 で関連付いた Wikipedia 記事のエントリ名をクエリとして様々なウェブコンテンツを検索し、その結果を関連付ける。本デモンストレーションでは、以下の API を利用して関連付けを行なったウェブコンテンツに対してリンクを張った。

- 日テレ API^(注2)
- Yahoo! ニュース・トピックス API^(注3)
- Google AJAX Search API^(注4)

(注1): <http://ja.wikipedia.org/>

(注2): <http://appli.ntv.co.jp/appli/api/news/index.html>

(注3): <http://public.news.yahoo.co.jp/api/>

(注4): <http://code.google.com/intl/ja/apis/ajaxsearch/>

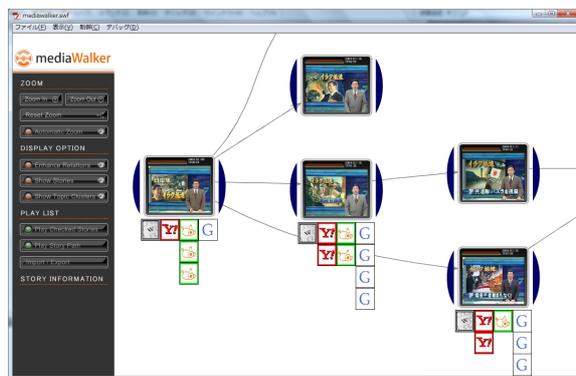


図1 mediaWalker II インタフェース。ニュースの時系列意味構造「トピックスレッド」に沿って、トピックの変遷を追跡できる。各ストーリーでは、映像を再生するとともに、関連ウェブコンテンツを閲覧できる。

3. mediaWalker II インタフェース

得られた関連ウェブコンテンツへのリンクを、mediaWalker インタフェースに追加した (図1)。これにより、特定のニュース番組から得られる情報だけでなく、関連する様々なウェブコンテンツを参照しながら、多角的な内容理解を促進できると考える。

4. おわりに

本デモンストレーションでは、ニュース映像アーカイブ中の映像を、時系列意味構造に基づいてストーリー単位で閲覧しながら、同時にウェブコンテンツを閲覧するために開発した mediaWalker II インタフェースを紹介した。今後は、アーカイブ中の他放送局のニュース映像との関連付け [1] 結果の利用や、放送映像とウェブコンテンツを総合的に利用した情報再編纂などに取り組む。

謝辞 本研究の一部は (財) 放送文化基金、科学研究費補助金及び国立情報学研究所との共同研究に基づく。

文 献

- [1] A. Ogawa, T. Takahashi, I. Ide, and H. Murase, "Cross-lingual retrieval of identical news events using image information," Proc. 14th Int. Multimedia Modeling Conf., Lecture Notes in Computer Science, Springer-Verlag, vol.4903, pp.287-296, Jan. 2008.
- [2] 井手一郎, 木下智義, 高橋友和, 佐藤真一, 村瀬 洋, "mediaWalker: 時系列意味構造に基づく映像アーカイブ探索インタフェース," 画像の認識・理解シンポジウム 2008 論文集, no.DS-5, pp.1670-1671, July 2008.
- [3] 井手一郎, 木下智義, 高橋友和, 孟 洋, 片山紀生, 佐藤真一, 村瀬 洋, "大量ニュース映像を対象とした時系列意味構造に基づく情報編纂手法の提案," 人工知能学論, vol.23, no.5, pp.282-292, Sept. 2008.
- [4] 井手一郎, 木下智義, 高橋友和, 佐藤真一, 村瀬 洋, "アーカイブされたニュース映像閲覧・編集インタフェース: mediaWalker," 画像の認識・理解シンポジウム 2009 論文集, no.DS-1, pp.1859-1860, July 2009.
- [5] T. Okuoka, T. Takahashi, D. Deguchi, I. Ide, and H. Murase, "Labeling news topic threads with Wikipedia entries," Proc. 11th IEEE Int. Symposium on Multimedia, pp.501-504, Dec. 2009.